EEC 118 Lecture #2: MOSFET Structure and Basic Operation

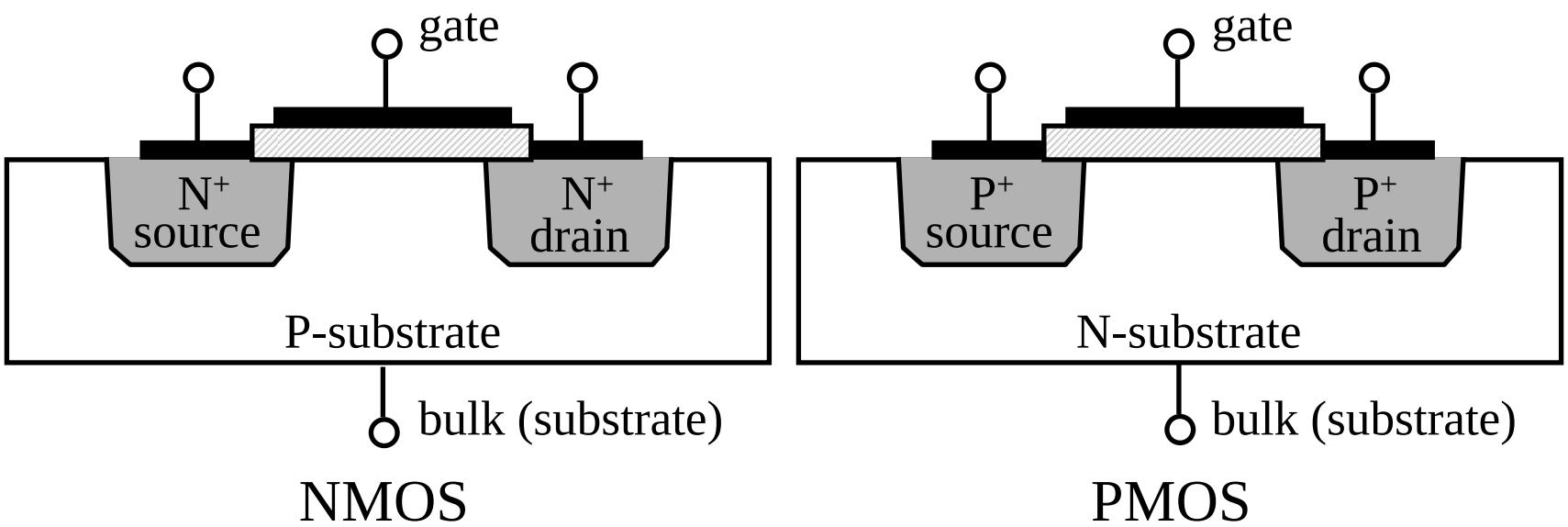
Rajeevan Amirtharajah
University of California, Davis
Jeff Parkhurst
Intel Corporation

Outline

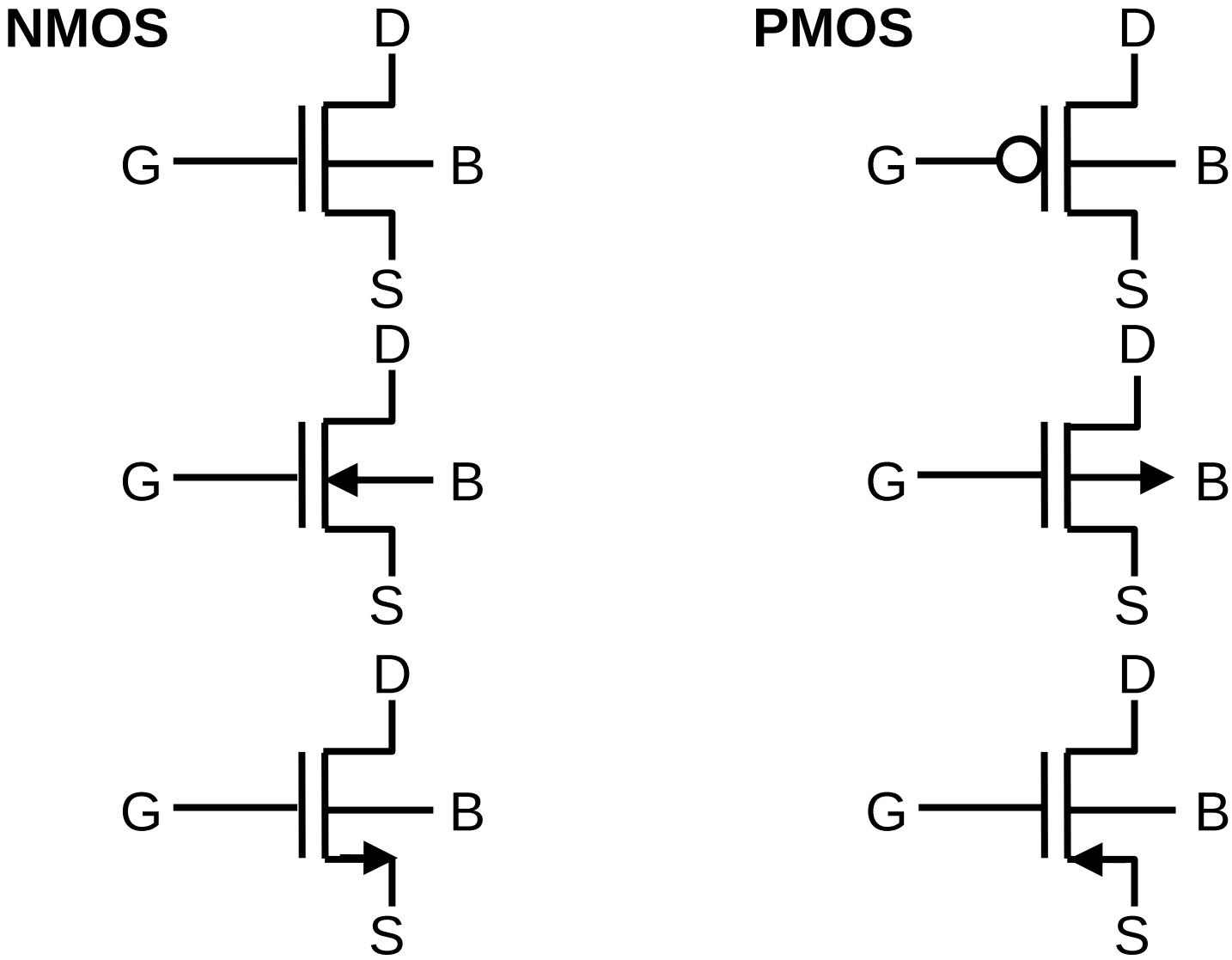
- **Finish Lecture 1 Slides**
- **Switch Example**
- **MOSFET Structure**
- **MOSFET Regimes of Operation**
- **Scaling**
- **Parasitic Capacitances**

MOS Transistor Types

- **Rabaey Ch. 3 (Kang & Leblebici Ch. 3)**
- **Two transistor types (analogous to bipolar NPN, PNP)**
  - NMOS: p-type substrate, n^+ source/drain, electrons are charge carriers
  - PMOS: n-type substrate, p^+ source/drain, holes are charge carriers

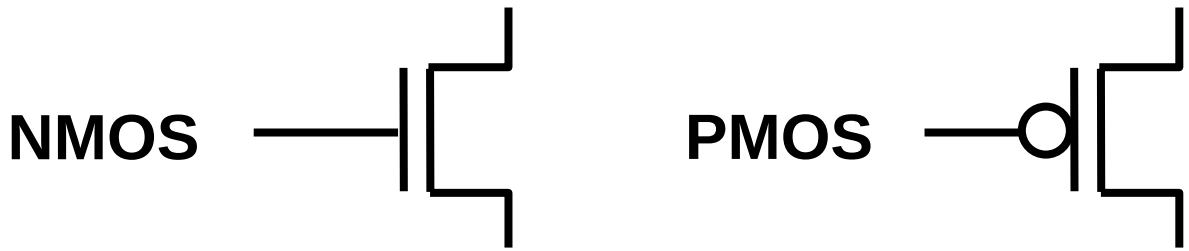


MOS Transistor Symbols



Note on MOS Transistor Symbols

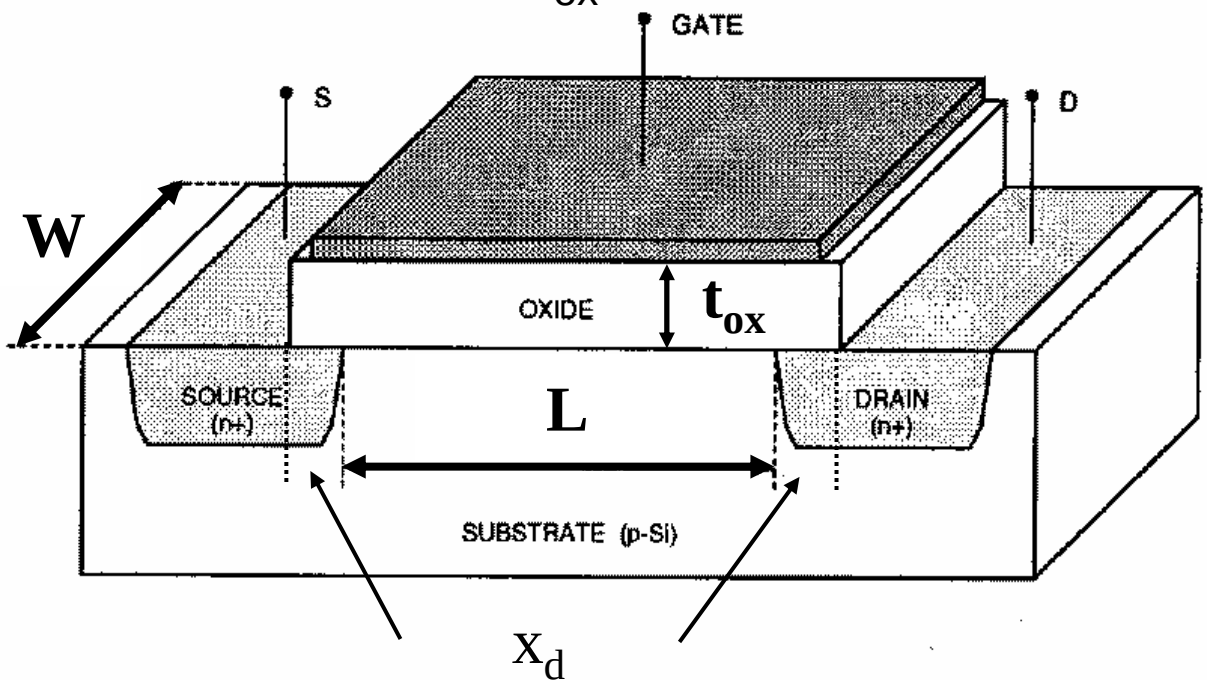
- **All symbols appear in literature**
  - Symbols with arrows are conventional in analog papers
  - PMOS with a bubble on the gate is conventional in digital circuits papers
- **Sometimes bulk terminal is ignored – implicitly connected to supply:**



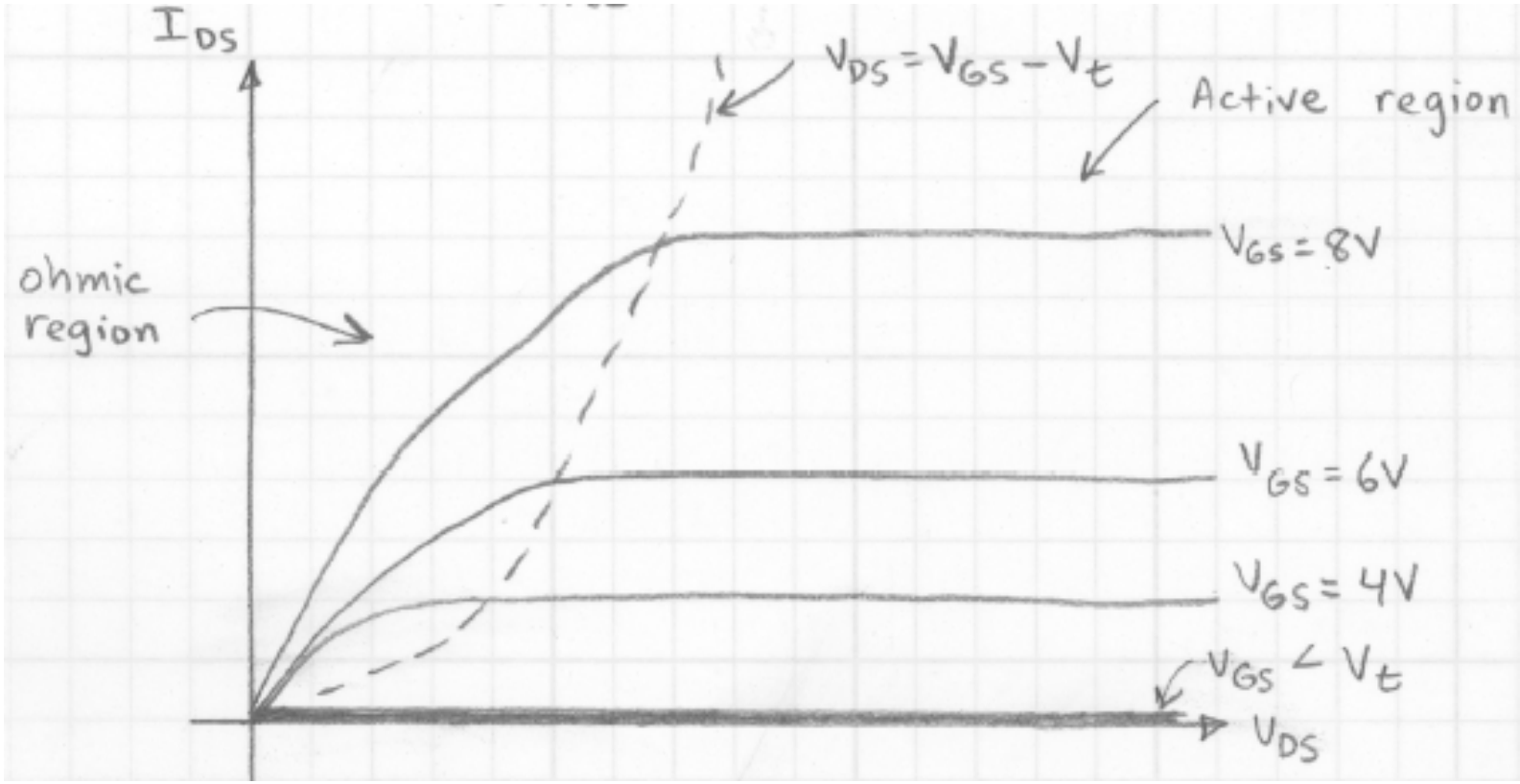
- **Unlike physical bipolar devices, source and drain are usually symmetric**

MOS Transistor Structure

- **Important transistor physical characteristics**
  - Channel length $L = L_D - 2x_d$ (K&L $L = Lgate - 2L_D$)
  - Channel width W
  - Thickness of oxide t_{ox}

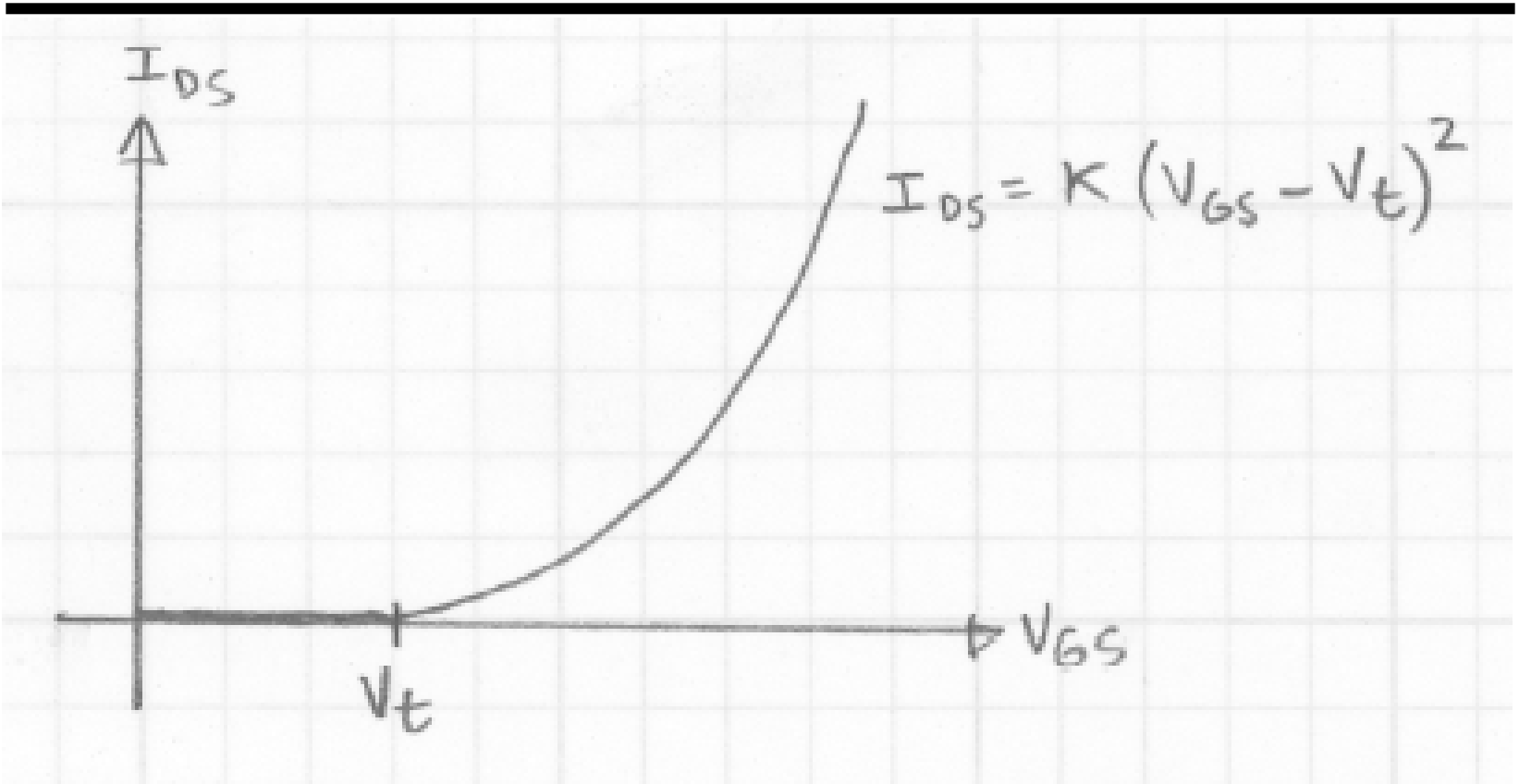


NMOS Transistor I-V Characteristics I



- **I-V curve vaguely resembles bipolar transistor curves**
  - Quantitatively very different
  - Turn-on voltage called Threshold Voltage V_T

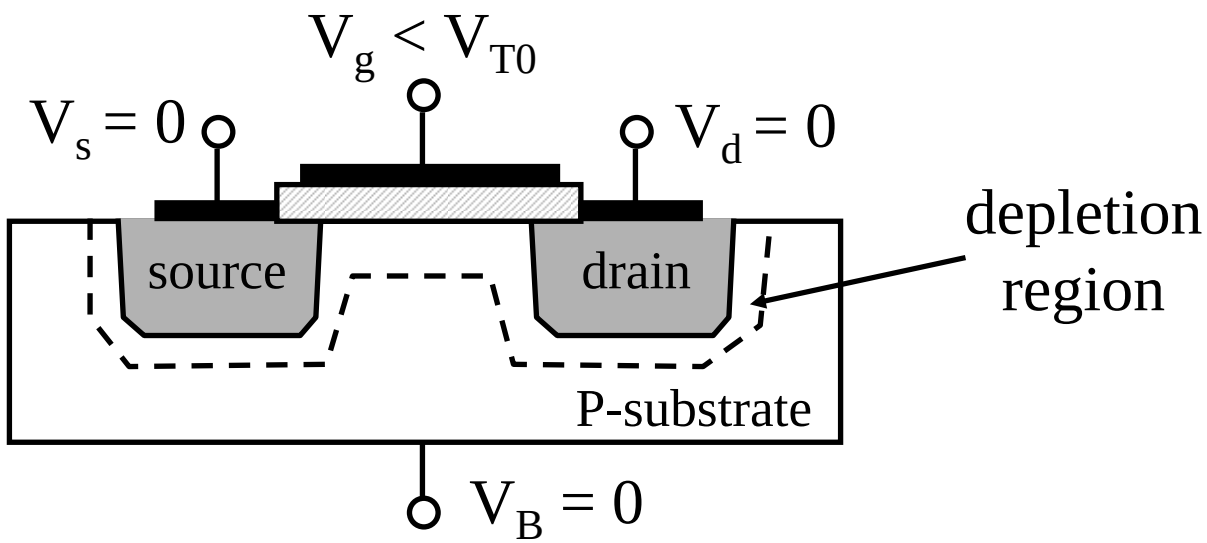
NMOS Transistor I-V Characteristics II



- **Drain current varies quadratically with gate-source voltage V_{GS} (in Saturation)**

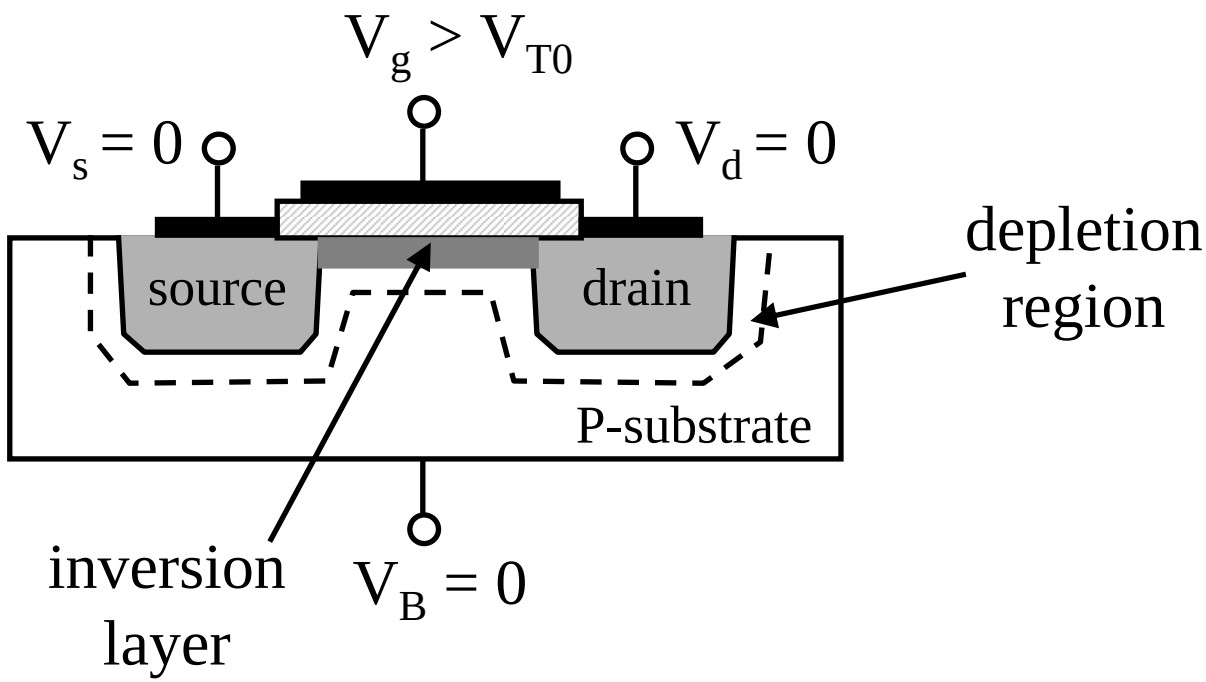
MOS Transistor Operation: Cutoff

- **Simple case: $V_D = V_S = V_B = 0$**
  - Operates as MOS capacitor (Cg = gate to channel)
  - Transistor in cutoff region
- **When $V_{GS} < V_{T0}$, depletion region forms**
  - No carriers in channel to connect S and D (Cutoff)



MOS Transistor Operation: Inversion

- **When $V_{GS} > V_{T0}$, inversion layer forms**
- **Source and drain connected by conducting n-type layer (for NMOS)**
  - Conducting p-type layer in PMOS



Threshold Voltage Components

- **Four physical components of the threshold voltage**

1. **Work function difference between gate and channel (depends on metal or polysilicon gate): Φ_{GC}**
2. **Gate voltage to invert surface potential: $-2\Phi_F$**
3. **Gate voltage to offset depletion region charge: Q_B/C_{ox}**
4. **Gate voltage to offset fixed charges in the gate oxide and oxide-channel interface: Q_{ox}/C_{ox}**

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$$ **: gate oxide capacitance per unit area**

Threshold Voltage Summary

- **If $V_{SB} = 0$ (no substrate bias):**

$$V_{T0} = \Phi_{GC} - 2\phi_F - \frac{Q_{B0}}{C_{ox}} - \frac{Q_{ox}}{C_{ox}} \quad \text{(K\&L 3.20)}$$

- **If $V_{SB} \neq 0$ (non-zero substrate bias)**

$$V_T = V_{T0} + \gamma\left(\sqrt{\left|-2\phi_F + V_{SB}\right|} - \sqrt{\left|2\phi_F\right|}\right) \quad (3.19)$$

- **Body effect (substrate-bias) coefficient:**

$$\gamma = \frac{\sqrt{2qN_A\varepsilon_{Si}}}{C_{ox}} \quad \text{(K\&L 3.24)}$$

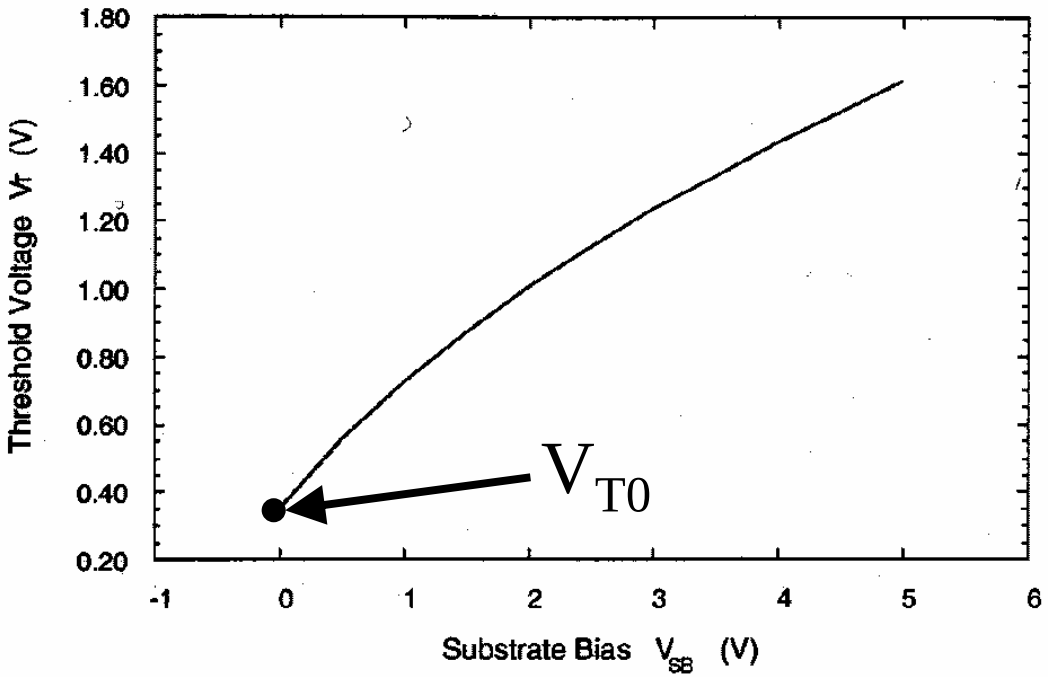
- **Threshold voltage increases as V_{SB} increases!**

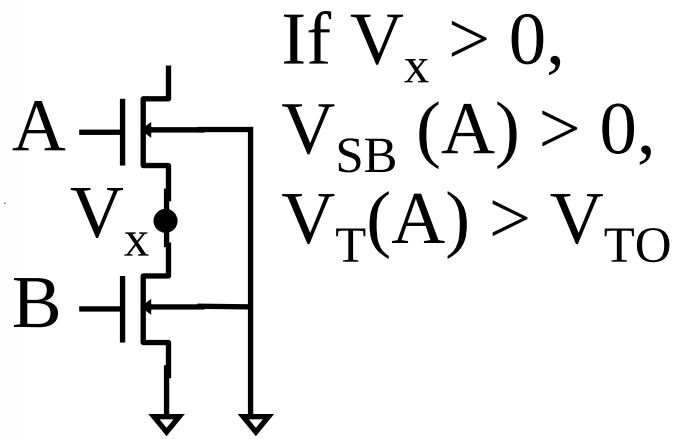
Threshold Voltage (NMOS vs. PMOS)

	NMOS	PMOS
Substrate Fermi potential	$\phi_F < 0$	$\phi_F > 0$
Depletion charge density	$Q_B < 0$	$Q_B > 0$
Substrate bias coefficient	$\gamma > 0$	$\gamma < 0$
Substrate bias voltage	$V_{SB} > 0$	$V_{SB} < 0$

Body Effect

- **Body effect: Source-bulk voltage V_{SB} affects threshold voltage of transistor**
  - Body normally connected to ground for NMOS, Vdd (Vcc) for PMOS
  - Raising source voltage increases V_T of transistor
  - Implications on circuit design: series stacks of devices





MOS Transistor Regions of Operation

- **Three main regions of operation**
- **<u>Cutoff</u>: $V_{GS} < V_T$**
  No inversion layer formed, drain and source are isolated by depleted channel. $I_{DS} \approx 0$
- **<u>Linear (Triode, Ohmic)</u>: $V_{GS} > V_T$, $V_{DS} < V_{GS}-V_T$**
  Inversion layer connects drain and source. Current is almost linear with V_{DS} (like a resistor)
- **<u>Saturation</u>: $V_{GS} > V_T$, $V_{DS} \geq V_{GS}-V_T$**
  Channel is "pinched-off". Current saturates (becomes independent of V_{DS}, to first order).

MOSFET Drain Current Overview

Saturation: $$I_D = \frac{\mu C_{ox}}{2}\frac{W}{L}\left(V_{GS} - V_T\right)^2\left(1 + \lambda V_{DS}\right)$$

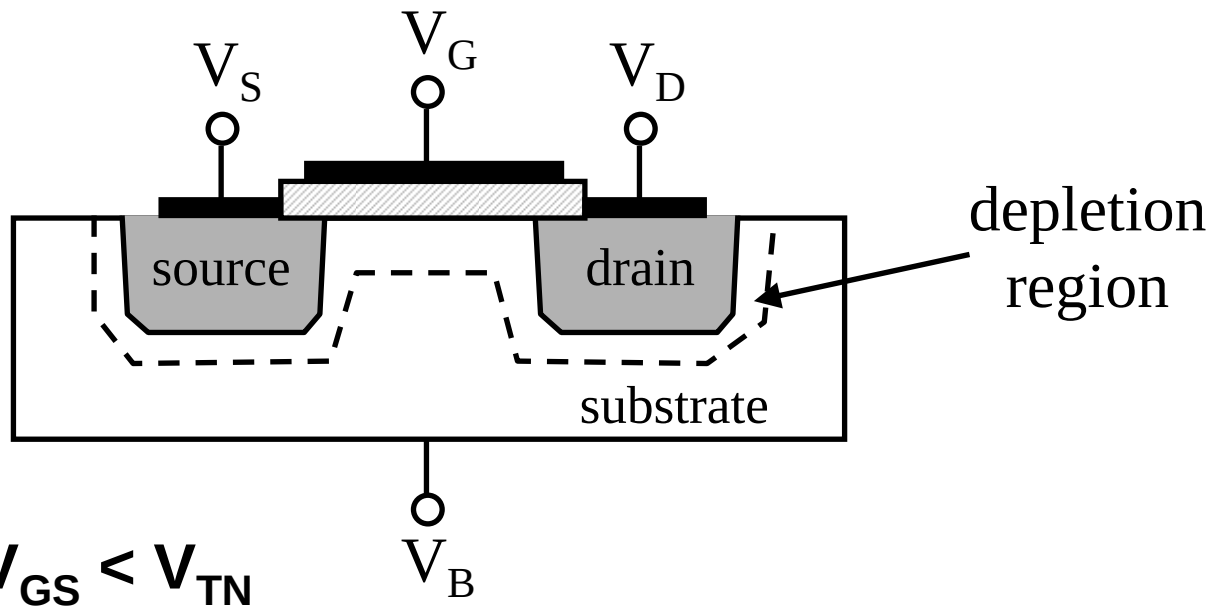
Linear (Triode, Ohmic):

$$I_D = \mu C_{ox}\frac{W}{L}\left(\left(V_{GS} - V_T\right)V_{DS} - \frac{V_{DS}^2}{2}\right)$$

Cutoff: $$I_D \approx 0$$

"Classical" MOSFET model, will discuss deep submicron modifications as necessary (Rabaey, Eqs. 3.25, 3.29)

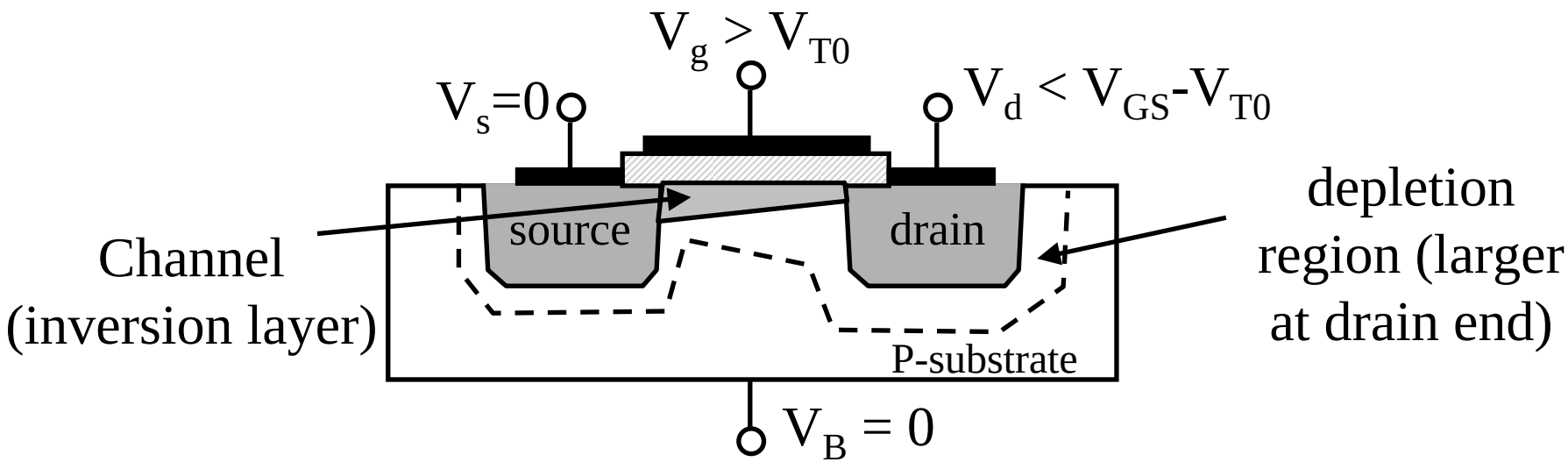
Cutoff Region



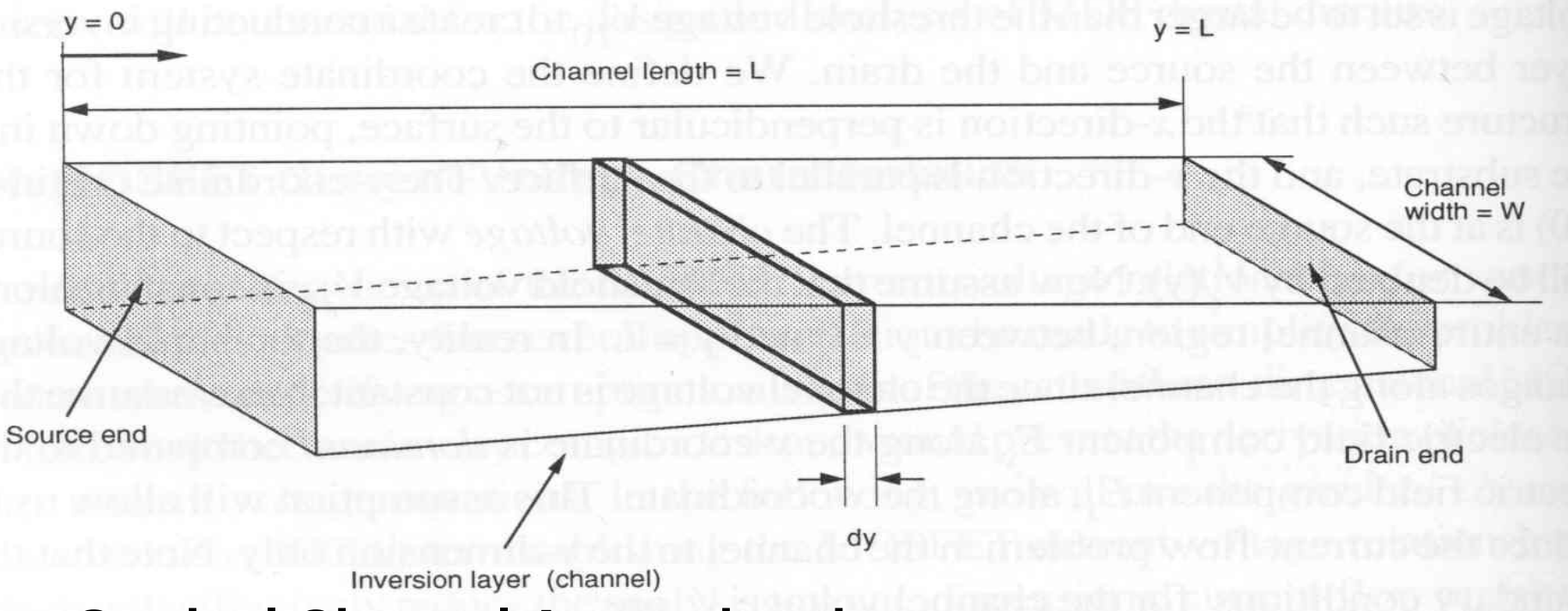
- **For NMOS: $V_{GS} < V_{TN}$**
- **For PMOS: $V_{GS} > V_{TP}$**
- **Depletion region – no inversion**
- **Current between drain and source is 0**
  - Actually there is always some leakage (subthreshold) current

Linear Region

- **When $V_{GS}>V_T$, an inversion layer forms between drain and source**
- **Current I_{DS} flows from drain to source (electrons travel from source to drain)**
- **Depth of channel depends on V between gate and channel**
  - Drain end narrower due to larger drain voltage
  - Drain end depth reduces as V_{DS} is increased



Linear Region I/V Equation Derivation



- **Gradual Channel Approximation:**
  - Assume dominant electric field in y-direction
  - Current is constant along channel
- **Integrate differential voltage drop $dV_c = I_D dR$ along y**

Linear Region I/V Equation

- **Valid for continuous channel from Source to Drain**

$$I_D = \mu_n C_{ox} \frac{W}{L}\left[(V_{GS} - V_T)V_{DS} - \tfrac{1}{2}V_{DS}^2\right]$$

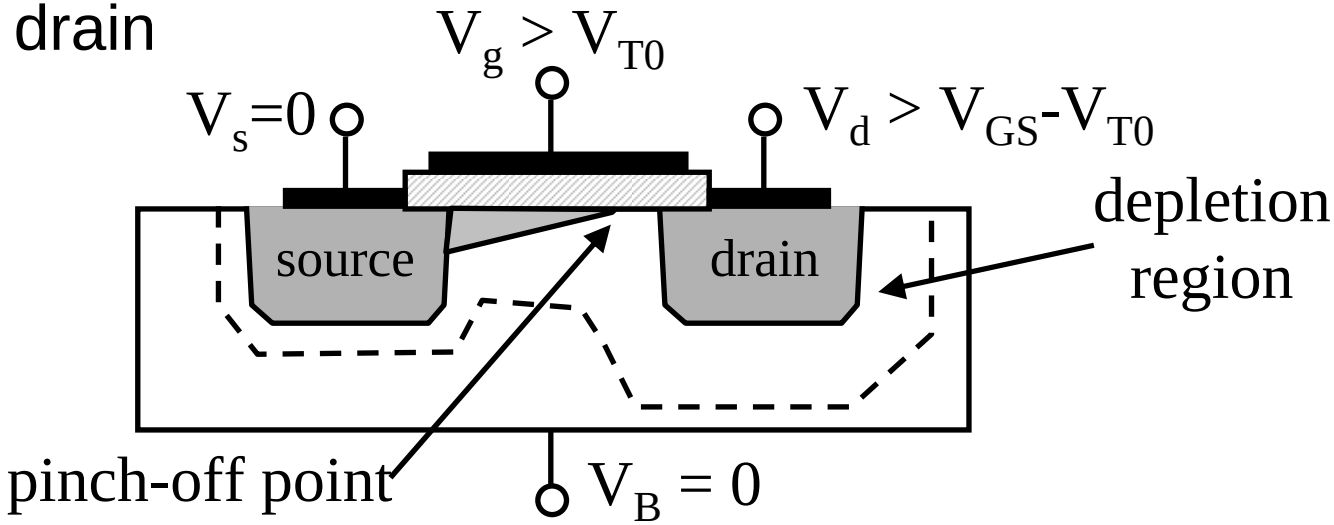
Device transconductance: $k_n = \mu_n C_{ox} \frac{W}{L}$

Process transconductance: $k_n' = \mu_n C_{ox}$

$$I_D = k_n' \frac{W}{L}\left[(V_{GS} - V_T)V_{DS} - \tfrac{1}{2}V_{DS}^2\right]$$

Saturation Region

- **When $V_{DS} = V_{GS} - V_T$:**
  - No longer voltage drop of V_T from gate to substrate at drain
  - Channel is “pinched off”
- **If V_{DS} is further increased, no increase in current I_{DS}**
  - As V_{DS} increased, pinch-off point moves closer to source
  - Channel between that point and drain is depleted
  - High electric field in depleted region accelerates electrons towards drain



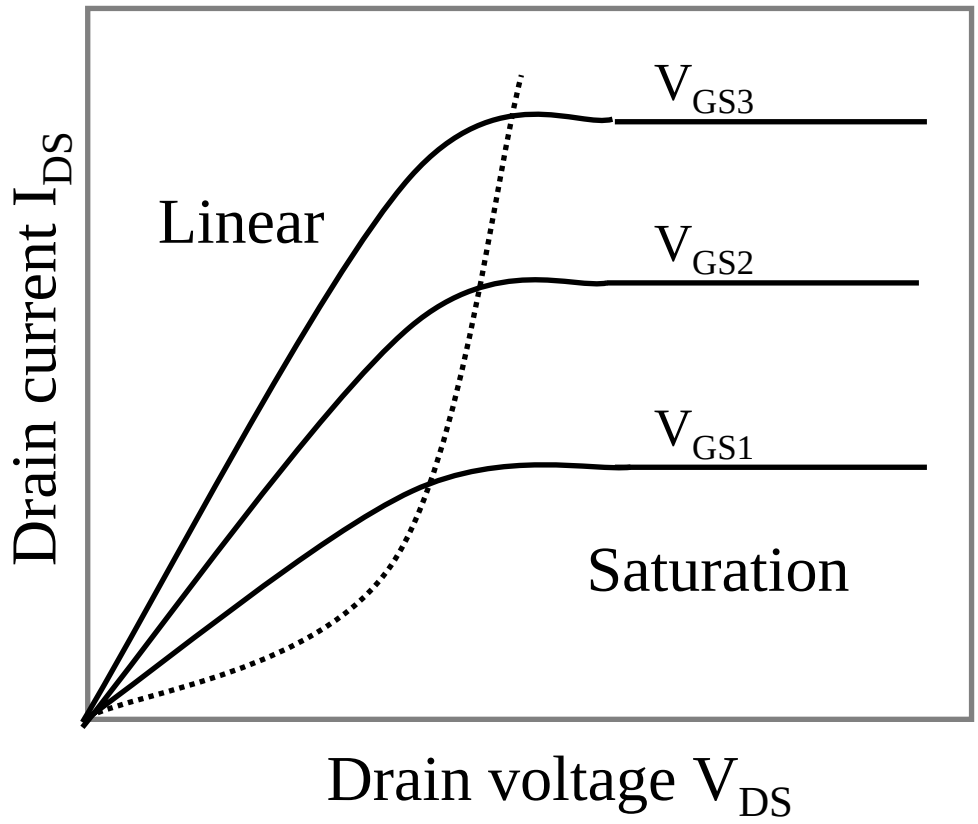
Saturation I/V Equation

- **As drain voltage increases, channel remains pinched off**
  - Channel voltage remains constant
  - Current saturates (no increase with increasing V_{DS})
- **To get saturation current, use linear equation with $V_{DS} = V_{GS} - V_T$**

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left(V_{GS} - V_{TN}\right)^2$$

MOS I/V Characteristics

- **I/V curve for ideal MOS device**
- **$V_{GS3} > V_{GS2} > V_{GS1}$**



Channel Length Modulation

- **In saturation, pinch-off point moves**
  - As V_{DS} is increased, pinch-off point moves closer to source
  - Effective channel length becomes shorter
  - Current increases due to shorter channel

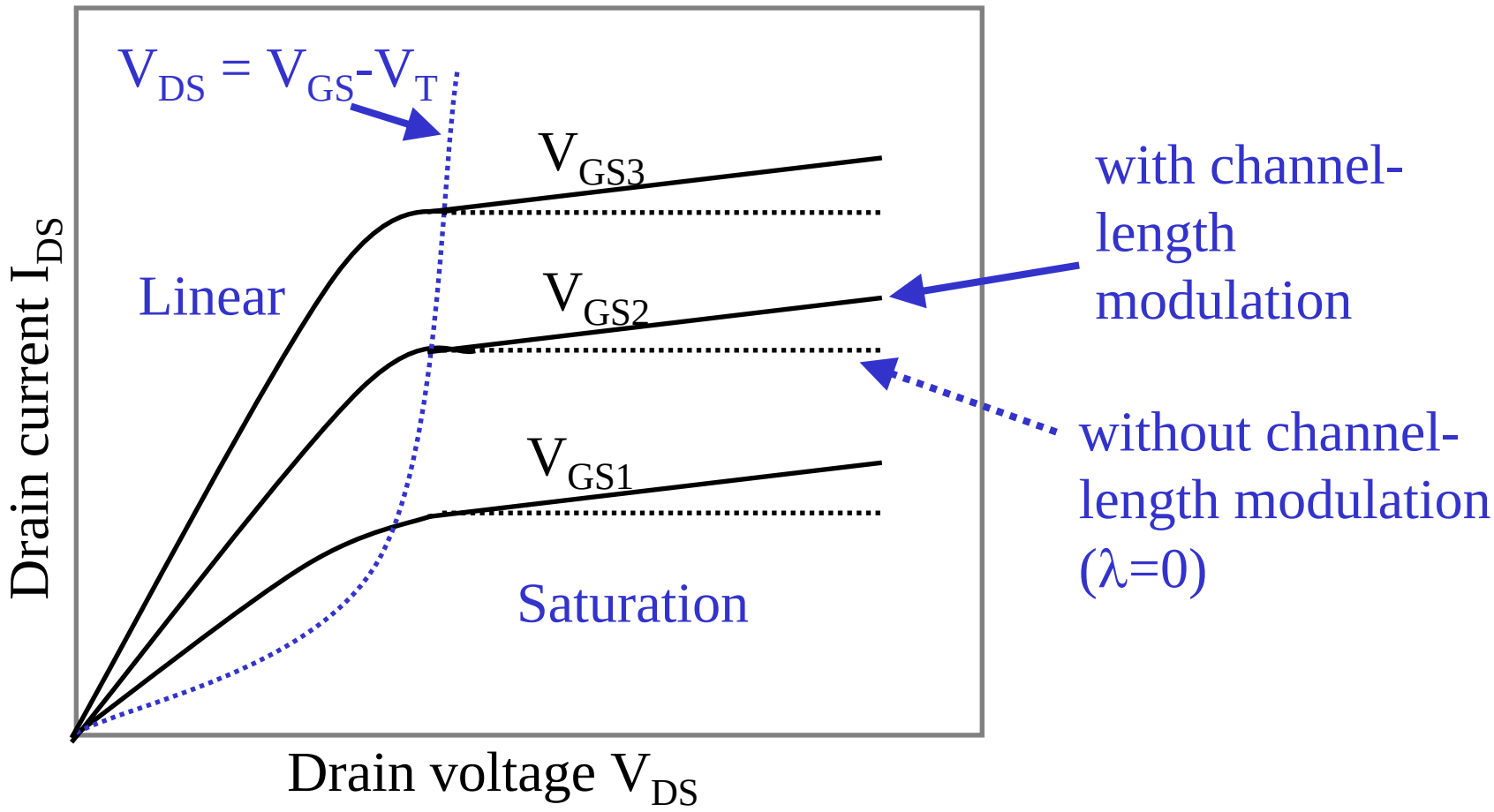
$$L' = L - \Delta L$$

$$I_D = \tfrac{1}{2}\mu_n C_{ox}\frac{W}{L}\left(V_{GS} - V_{TN}\right)^2\left(1 + \lambda V_{DS}\right)$$

λ = channel length modulation coefficient

MOS I/V Curve Summary

I/V curve for non-ideal NMOS device:



MOS I/V Equations Summary

Cutoff

$$\begin{aligned} V_{GS} &< V_{TN} \\ V_{GS} &> V_{TP} \end{aligned} \Rightarrow I_D = 0$$

Linear

$$\begin{aligned} V_{GS} \geq V_{TN}, &\quad V_{DS} < V_{GS} - V_{TN} \\ V_{GS} \leq V_{TP}, &\quad V_{DS} > V_{GS} - V_{TP} \end{aligned} \Rightarrow I_D = \mu C_{ox} \frac{W}{L} \left[\left(V_{GS} - V_T \right) V_{DS} - \tfrac{1}{2} V_{DS}^2 \right]$$

Saturation

$$\begin{aligned} V_{GS} \geq V_{TN}, &\quad V_{DS} \geq V_{GS} - V_{TN} \\ V_{GS} \leq V_{TP}, &\quad V_{DS} \leq V_{GS} - V_{TP} \end{aligned} \Rightarrow I_D = \tfrac{1}{2} \mu C_{ox} \frac{W}{L} \left(V_{GS} - V_T \right)^2 \left(1 + \lambda V_{DS} \right)$$

Note: if $V_{SB} \neq 0$, need to recalculate V_T from V_{T0}

A Fourth Region: Subthreshold

Subthreshold: $$I_D = I_S e^{\frac{V_{GS}}{n\,kT/q}} \left(1 - e^{-\frac{V_{DS}}{kT/q}} \right)$$

- **Sometimes called “weak inversion” region**
- **When V_{GS} near V_T, drain current has an exponential dependence on gate to source voltage**
  - Similar to a bipolar device
- **Not typically used in digital circuits**
  - Sometimes used in very low power digital applications
  - Often used in low power analog circuits, e.g. quartz watches

MOSFET Scaling Effects

- **Rabaey Section 3.5 (Kang & Leblebici Section 3.5)**
- **Scaling provides enormous advantages**
  - Scale linear dimension (channel length) by factor S > 1
  - Better area density, yield, performance
- **Two types of scaling**
  - Constant field scaling (full scaling)
    - $A' = A/S^2$; $L' = L/S$; $W' = W/S$; $I_D' = I_D/S$; $P' = P/S^2$; $V_{dd}' = V_{dd}/S$
    - Power Density P'/A' = stays the same
  - Constant voltage scaling
    - $A' = A/S^2$; $L' = L/S$; $W' = W/S$; $I_D' = I_D*S$; $P' = P*S$; $V_{dd}' = V_{dd}$
    - Power Density P'/A' = S^3*P (Reliability issue)

Change these two

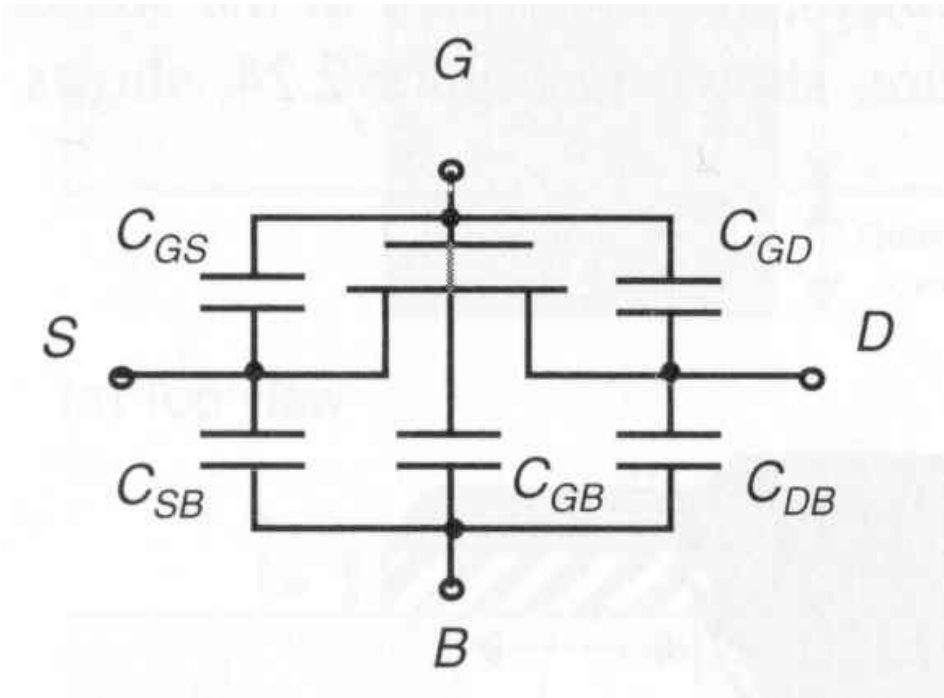
This changed as well

Short Channel Effects

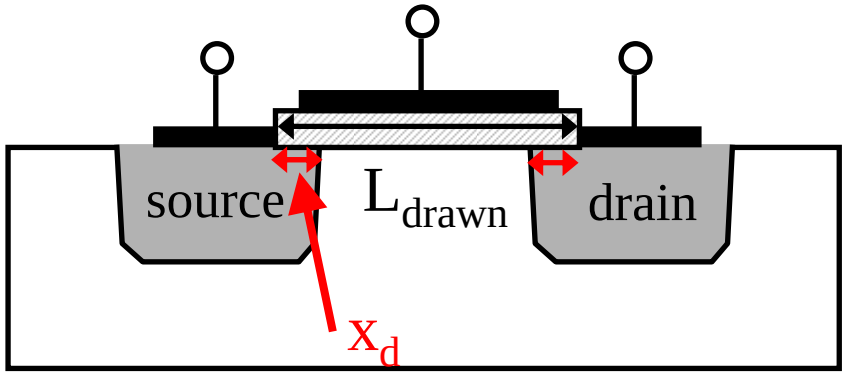
- **As geometries are scaled down**
  - V_T (effective) goes lower
  - Effective channel length decreases
  - Sub-threshold Ids occurs
    - Current goes from drain to source while Vgs < Vt
  - Tox is scaled which can cause reliability problems
    - Can't handle large Vg without hot electron effects
      - Changes the Vt when carriers imbed themselves in the oxide
  - Interconnects scale
    - Electromigration and ESD become issues

MOSFET Capacitances

- **Rabaey Section 3.3 (Kang & Leblebici Section 3.6)**
- **Oxide Capacitance**
  - Gate to Source overlap
  - Gate to Drain overlap
  - Gate to Channel
- **Junction Capacitance**
  - Source to Bulk junction
  - Drain to Bulk junction



Oxide Capacitances: Overlap

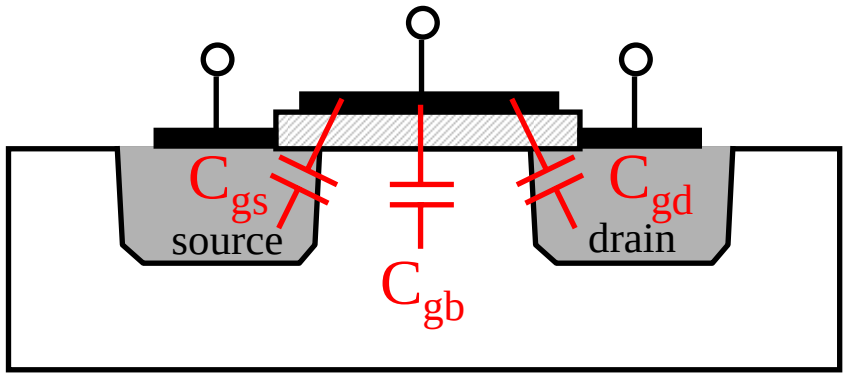


- **Overlap capacitances**
  - Gate electrode overlaps source and drain regions
  - x_d is overlap length on each side of channel
  - $L_{eff} = L_{drawn} - 2x_d$ (effective channel length)
  - Overlap capacitance:

$$C_{GSO} = C_{GDO} = C_{ox} W x_d$$ Assume x_d equal on both sides

Total Oxide Capacitance

- **Total capacitance consists of 2 components**
  - Overlap capacitance
  - Channel capacitance



- **Cutoff:**
  - No channel connecting to source or drain
  - $C_{GS} = C_{GD} = C_{ox}Wx_d$
  - $C_{GB} = C_{ox}WL_{eff}$
  - Total Gate Capacitance = $C_G = C_{ox}WL$

Oxide Capacitances: Channel

- **Linear mode**
  - Channel spans from source to drain
  - Channel Capacitance split equally between S and D

$$C_{GS} = \frac{1}{2} C_{ox} W L_{eff} \quad C_{GD} = \frac{1}{2} C_{ox} W L_{eff} \quad C_{GB} = 0$$

  - Total Gate capacitance $C_G = C_{ox} WL$

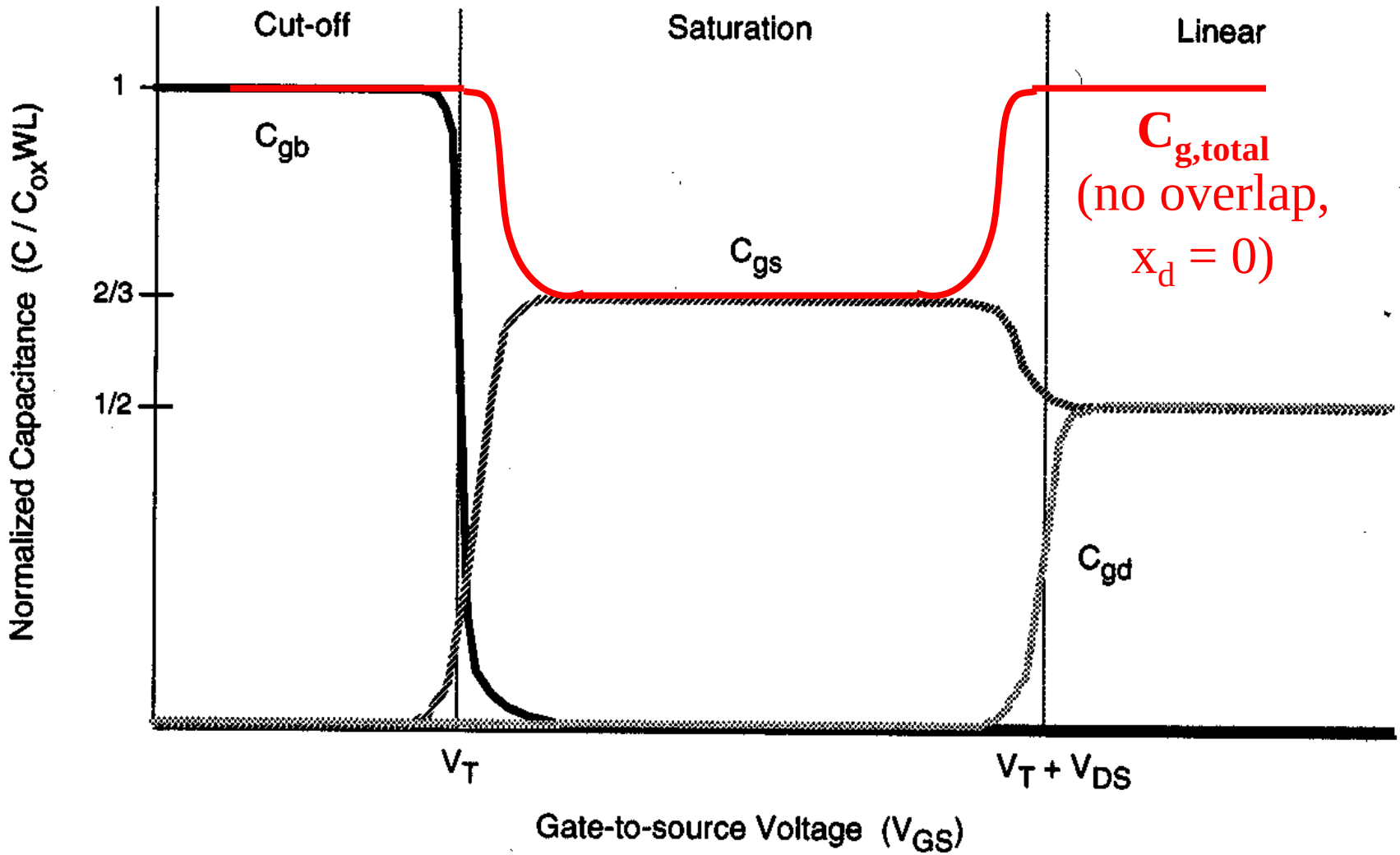
- **Saturation regime**
  - Channel is pinched off: Channel Capacitance --

$$C_{GD} = W x_d C_{ox} \quad C_{GS} = \frac{2}{3} C_{ox} W L_{eff} + C_{OX} W x_d \quad C_{GB} = 0$$

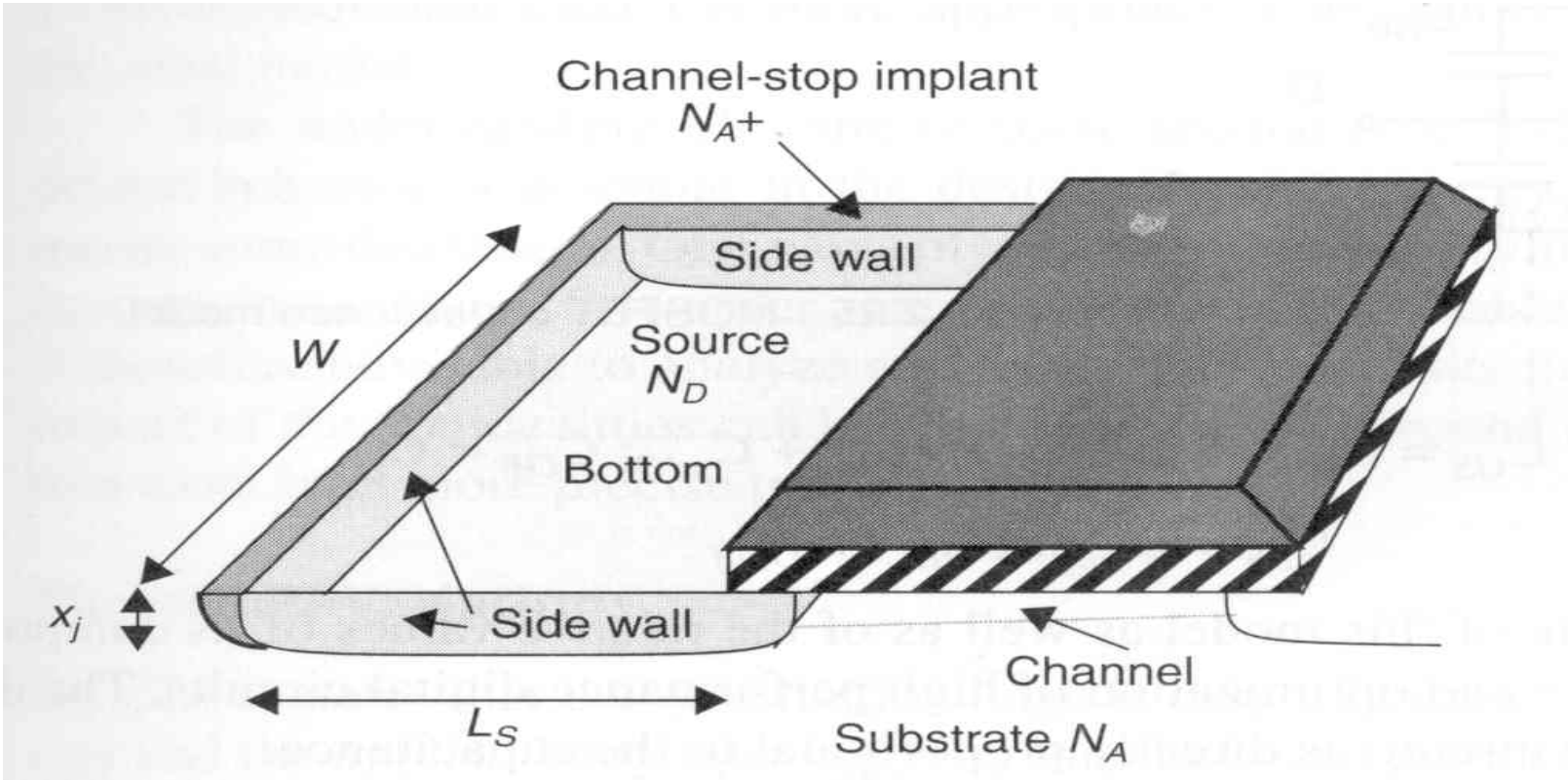
  - Total Gate capacitance:

$$C_G = 2/3\ C_{ox} W L_{eff} + 2 x_d W C_{OX}$$

Oxide Capacitances: Channel



Junction Capacitance



Reverse-biased P-N junctions!
Capacitance depends on reverse-bias voltage.

Junction Capacitance

For a P-N junction: $$C_j = \frac{A}{2}\sqrt{\frac{2q\varepsilon}{V_0 - V}\frac{N_d N_a}{N_d + N_a}}$$

If V=0, cap/area = $$C_{j0} = \sqrt{\frac{q\varepsilon_{Si}}{2V_0}\frac{N_d N_a}{N_d + N_a}}$$

General form: $$C_j = \frac{AC_{j0}}{\left(1 - \frac{V}{V_0}\right)^m}$$

m = grading coefficient (0.5 for abrupt junctions)
(0.3 for graded junctions)

Junction Capacitance

- **Junction with substrate**
  - Bottom area = W * L_S (length of drain/source)
  - Total cap = C_j
- **Junction with sidewalls**
  - "Channel-stop implant"
  - Perimeter = $2L_S + W$
  - Area = P * X_j
  - Total cap = C_{jsw}
- **Total junction cap C = $C_j + C_{jsw}$**

Junction Capacitance

- **Voltage Equivalence Factor**
  - Creates an average capacitance value for a voltage transition, defined as $\Delta Q/\Delta V$

$$C_{eq} = \frac{-AC_{j0}V_0}{(V_2 - V_1)(1-m)}\left(\left(1-\frac{V_2}{V_0}\right)^{1-m} - \left(1-\frac{V_1}{V_0}\right)^{1-m}\right) = AK_{eq}C_{j0}$$

$$K_{eq} = \frac{-2\sqrt{V_0}}{(V_2 - V_1)}\left(\sqrt{V_0 - V_2} - \sqrt{V_0 - V_1}\right)$$ (abrupt junction only)

$$C_{db} = AK_{eq}C_{j0} + PX_jK_{eqsw}C_{jsw0}$$

Example: Junction Cap

- **Consider the following NMOS device**
  - Substrate doping: $N_A = 10^{15}$ cm^{-3}
  - Source/drain doping: $N_D = 2 \times 10^{20}$ cm^{-3}
  - Channel-stop doping: 10X substrate doping
  - Drain length L_D = 1um
  - Transistor W = 10um
  - Junction depth Xj = 0.5um, abrupt junction

- **Find capacitance of drain-bulk junction when drain voltage = 3V**

Next Topic: Inverters

- **Inverter Characteristics**
  - Transfer functions, noise margins, resistive and nonlinear loads
- **CMOS Inverters**